Many calculations in calculus are easier to do using *differentials*. Furthermore, differentials and the related *differential forms* are often used in applications, especially (but not only) to physics. The official textbook covers differentials, but incompletely and only in one minor application. It then uses differentials again later (mostly in material for Calculus 2 and 3), but they are useful much earlier. So I will make heavy use of them.

Notation and terminology

If y is a variable quantity, then $\mathrm{d}y$ is the **differential** of y. You can think of $\mathrm{d}y$ as indicating an infinitely small (infinitesimal) change in the value of y, or the amount by which y changes when an infinitesimal change is made. A precise definition is at the end of these notes, but you will *not* be tested directly on that; what you need to know is how to *use* differentials.

Note that $\mathrm{d}y$ is *not* d times y, and d is also *not* exactly a function of y. Rather, y (being a *variable* quantity) should itself be a function of some other quantity x, and $\mathrm{d}y$ is also a function of a sort; so d is an *operator*: something that turns one function into another function. However, an expression like $A\,\mathrm{d}y$ does involve multiplication: it is A times the differential of y.

We often divide one differential by another; for example, $\mathrm{d}y/\mathrm{d}x$ is the result of dividing the differential of y by the differential of x. The textbook introduces this notation early to stand for the *derivative* of y with respect to x, and indeed it is that; but what the book doesn't tell you is that $\mathrm{d}y/\mathrm{d}x$ literally is $\mathrm{d}y$ divided by $\mathrm{d}x$. Unfortunately, $\mathrm{d}^2y/\mathrm{d}x^2$, the second derivative, is *not* literally $\mathrm{d}^2y = \mathrm{d}(\mathrm{d}y)$ divided by $\mathrm{d}x^2 = (\mathrm{d}x)^2$; for this reason, I prefer the notation $(\mathrm{d}/\mathrm{d}x)^2 y$, meaning $(\mathrm{d}/\mathrm{d}x)(\mathrm{d}/\mathrm{d}x)y = (\mathrm{d}/\mathrm{d}x)(\mathrm{d}y/\mathrm{d}x) = \mathrm{d}(\mathrm{d}y/\mathrm{d}x)/\mathrm{d}x$.

The Chain Rule

The most important fact about differentials is this: If f is a differentiable function, then

$$\mathrm{d}f(u) = f'(u)\,\mathrm{d}u.$$

That is, the differential of $f(u)$ equals $f'(u)$ times the differential of u, where f' is the derivative of f (as a function). This fact not only shows the relationship between differentials and derivatives, but also (because u could be any quantity) it encapsulates the **Chain Rule** in differential form. The Chain Rule is an important principle in calculus, which is often difficult to learn how to use; but with differentials it is easy.

In particular, if $y = f(x)$, then

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{f'(x)\,\mathrm{d}x}{\mathrm{d}x} = f'(x),$$

so $\mathrm{d}y$ divided by $\mathrm{d}x$ really is the derivative.

For example, suppose that you have discovered (say from the definition as a limit) that the derivative of $f(x) = x^2$ is $f'(x) = 2x$. Then this fact can be expressed in differential form:

$$\mathrm{d}(x^2) = \mathrm{d}f(x) = f'(x)\,\mathrm{d}x = 2x\,\mathrm{d}x. \qquad (*)$$

Conversely, if (by performing a calculation with differentials) you discover the equation (*) above, then you know the derivative of f as well:

$$f'(x) = \frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{\mathrm{d}(x^2)}{\mathrm{d}x} = \frac{2x\,\mathrm{d}x}{\mathrm{d}x} = 2x.$$

Whichever of these facts you discover first, once you know them, you know something even more general:

$$\mathrm{d}(u^2) = 2u\,\mathrm{d}u.$$

(The power to derive this from equation (*) is the Chain Rule.) The value of this is that u can be any expression whatsoever; for example, if $u = x^2$ again, then

$$\mathrm{d}(x^4) = \mathrm{d}\big((x^2)^2\big) = 2(x^2)\,\mathrm{d}(x^2) = 2x^2(2x\,\mathrm{d}x) = 4x^3\,\mathrm{d}x.$$

So now you have learnt a new derivative, without having to calculate it from scratch.

Rules of differentiation

Every theorem about derivatives of functions may also be expressed as a theorem about differentials. Here are the most common rules:

- The Constant Rule: $\mathrm{d}k = 0$ if k is constant.
- The Sum Rule: $\mathrm{d}(u + v) = \mathrm{d}u + \mathrm{d}v$.
- The Translate Rule: $\mathrm{d}(u + k) = \mathrm{d}u$ if k is constant.
- The Difference Rule: $\mathrm{d}(u - v) = \mathrm{d}u - \mathrm{d}v$.
- The Product Rule: $\mathrm{d}(uv) = v\,\mathrm{d}u + u\,\mathrm{d}v$.
- The Multiple Rule: $\mathrm{d}(ku) = k\,\mathrm{d}u$ if k is constant.
- The Quotient Rule: $\mathrm{d}\left(\frac{u}{v}\right) = \frac{v\,\mathrm{d}u - u\,\mathrm{d}v}{v^2}$.
- The Power Rule: $\mathrm{d}\big(u^k\big) = ku^{k-1}\,\mathrm{d}u$ if k is constant.
- The Root Rule: $\mathrm{d}(\sqrt[k]{u}) = \frac{\sqrt[k]{u}\,\mathrm{d}u}{ku}$ if k is constant.

Of these, only the Constant Rule, the Sum Rule, the Product Rule, and the Power Rule are absolutely necessary, since every other expression built out of the operations in the rules above can be built out of the operations in these four rules. However, it is often handy to use all of these rules, even the Root Rule (which is not in the textbook). It is up to you how many of these rules to learn.

In addition, every time that you learn the derivative of a new function, you learn a new rule for differentials, by applying the Chain Rule to that function. We have already seen an example of this: applying the Chain Rule to the function $f(x) = x^2$ gives the special case of the Power Rule for $k = 2$. Here are a few other functions whose derivatives you will learn, expressed as rules for differentials:

- $\mathrm{d}(\mathrm{e}^u) = \mathrm{e}^u\,\mathrm{d}u$.
- $\mathrm{d}(\ln u) = \frac{\mathrm{d}u}{u}$.
- $\mathrm{d}(\sin u) = \cos u\,\mathrm{d}u$.
- $\mathrm{d}(\cos u) = -\sin u\,\mathrm{d}u$.
- $\mathrm{d}(\arctan u) = \frac{\mathrm{d}u}{u^2 + 1}$.

And more!

Notice that every one of these rules turns the differential on the left into a sum of terms (possibly only one term, or none in the case of the Constant Rule), each of which is an ordinary expression multiplied by a differential (or something algebraically equivalent to this). If, when you are calculating the differential of an expression, your result at any stage is *not* like this, then you have made a mistake!

Using differentials

The main technique for using differentials is simply to take the differential of both sides of an equation. However, you may only do this to an equation that holds *generally*, but *not* to an equation that holds only for *particular* values of the variables. (Ultimately, this is because d is an operator, not a function, so it must be applied to entire functions, not only to particular values of those functions.)

The simplest case is an equation such as $y = \mathrm{e}^{x^2}$, when we want the derivative of y with respect to x. So:

$$\begin{aligned} y &= \mathrm{e}^{x^2}; \\ \mathrm{d}y &= \mathrm{d}\big(\mathrm{e}^{x^2}\big) = \mathrm{e}^{x^2}\,\mathrm{d}(x^2) = \mathrm{e}^{x^2} \cdot 2x\,\mathrm{d}x = 2x\mathrm{e}^{x^2}\,\mathrm{d}x; \\ \frac{\mathrm{d}y}{\mathrm{d}x} &= 2x\mathrm{e}^{x^2}. \end{aligned}$$

Now we have the derivative. If we want the second derivative, then we do this again:

$$
\begin{aligned}
\mathrm{d}y/\mathrm{d}x &= 2x\mathrm{e}^{x^2}; \\
\mathrm{d}(\mathrm{d}y/\mathrm{d}x) &= \mathrm{d}\left(2x\mathrm{e}^{x^2}\right) = \mathrm{e}^{x^2}\,\mathrm{d}(2x) + 2x\,\mathrm{d}\left(\mathrm{e}^{x^2}\right) \\
&= \mathrm{e}^{x^2}\cdot 2\,\mathrm{d}x + 2x\cdot 2x\mathrm{e}^{x^2}\,\mathrm{d}x = \left(2\mathrm{e}^{x^2} + 4x^2\mathrm{e}^{x^2}\right)\mathrm{d}x; \\
(\mathrm{d}/\mathrm{d}x)^2 y = \frac{\mathrm{d}(\mathrm{d}y/\mathrm{d}x)}{\mathrm{d}x} &= 2\mathrm{e}^{x^2} + 4x^2\mathrm{e}^{x^2}.
\end{aligned}
$$

Now we have the second derivative (also written $\mathrm{d}^2y/\mathrm{d}x^2$).

The previous example began with an equation solved for y. But we don't need this; suppose instead that we have $y^5 + x^2 = x^5 + y$ (which *cannot* be solved for either variable using the usual algebraic operations of addition, subtraction, multiplication, division, powers, and roots). Undaunted, we forge ahead anyway:

$$
\begin{aligned}
y^5 + x^2 &= x^5 + y; \\
\mathrm{d}(y^5 + x^2) &= \mathrm{d}(x^5 + y); \\
\mathrm{d}(y^5) + \mathrm{d}(x^2) &= \mathrm{d}(x^5) + \mathrm{d}y; \\
5y^{5-1}\,\mathrm{d}y + 2x^{2-1}\,\mathrm{d}x &= 5x^{5-1}\,\mathrm{d}x + \mathrm{d}y; \\
5y^4\,\mathrm{d}y - \mathrm{d}y &= 5x^4\,\mathrm{d}x - 2x\,\mathrm{d}x; \\
(5y^4 - 1)\,\mathrm{d}y &= (5x^4 - 2x)\,\mathrm{d}x; \\
\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{5x^4 - 2x}{5y^4 - 1}.
\end{aligned}
$$

This process is called **implicit differentiation**.

The second derivative is a little simpler at first (or it would be if we didn't have to use the Quotient Rule!), but there is a twist at the end:

$$
\begin{aligned}
\mathrm{d}y/\mathrm{d}x &= \frac{5x^4 - 2x}{5y^4 - 1}; \\
\mathrm{d}(\mathrm{d}y/\mathrm{d}x) &= \mathrm{d}\left(\frac{5x^4 - 2x}{5y^4 - 1}\right) = \frac{(5y^4 - 1)\,\mathrm{d}(5x^4 - 2x) - (5x^4 - 2x)\,\mathrm{d}(5y^4 - 1)}{(5y^4 - 1)^2} \\
&= \frac{(5y^4 - 1)(20x^3 - 2)\,\mathrm{d}x - (5x^4 - 2x)(20y^3)\,\mathrm{d}y}{(5y^4 - 1)^2} \\
&= \frac{20x^3 - 2}{5y^4 - 1}\,\mathrm{d}x - \frac{20y^3(5x^4 - 2x)}{(5y^4 - 1)^2}\,\mathrm{d}y; \\
(\mathrm{d}/\mathrm{d}x)^2 y = \frac{\mathrm{d}(\mathrm{d}y/\mathrm{d}x)}{\mathrm{d}x} &= \frac{20x^3 - 2}{5y^4 - 1} - \frac{20y^3(5x^4 - 2x)}{(5y^4 - 1)^2}\,\frac{\mathrm{d}y}{\mathrm{d}x} \\
&= \frac{20x^3 - 2}{5y^4 - 1} - \frac{20y^3(5x^4 - 2x)}{(5y^4 - 1)^2}\,\frac{5x^4 - 2x}{5y^4 - 1}
\end{aligned}
$$

(which could be simplified further). Notice that I substitute the known expression for $\mathrm{d}y/\mathrm{d}x$ in the last step.

Another handy application of differentials is the case where both quantities x and y may be expressed as functions of some other quantity t. If we start with the same equation as above, then this will give us an equation relating the derivatives with respect to t:

$$\begin{aligned} y^5 + x^2 &= x^5 + y; \\ \mathrm{d}(y^5 + x^2) &= \mathrm{d}(x^5 + y); \\ \mathrm{d}(y^5) + \mathrm{d}(x^2) &= \mathrm{d}(x^5) + \mathrm{d}y; \\ 5y^{5-1}\,\mathrm{d}y + 2x^{2-1}\,\mathrm{d}x &= 5x^{5-1}\,\mathrm{d}x + \mathrm{d}y; \\ 5y^4\frac{\mathrm{d}y}{\mathrm{d}t} + 2x\frac{\mathrm{d}x}{\mathrm{d}t} &= 5x^4\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\mathrm{d}y}{\mathrm{d}t}. \end{aligned}$$

If we have information about one or both of these derivatives, then this equation will often give us useful information to solve a problem. This situation is called **related rates**, since derivatives can be viewed as rates of change (especially derivatives with respect to time t, although the t in the equation above doesn't have to be time).

Appendix: Definitions and proofs

The operator d is applied directly to a *variable quantity* rather than to a function in the usual sense. That is, we talk about $\mathrm{d}y$ rather than $\mathrm{d}f$, and while we do use $\mathrm{d}f(x)$, that is just because $f(x)$ is a variable quantity if x is. That is, $\mathrm{d}f(x)$ means $\mathrm{d}\big(f(x)\big)$.

So first, I need to formalize the concept of a variable quantity. If we choose a specific variable, which we call the independent variable, that every variable quantity can be expressed as a function of, then we can formalize each quantity as its corresponding function. But the choice of independent variable can be fairly arbitrary, and it's not always clear that one always exists either.

So, let us instead formalize a variable quantity y as an operation that, given a variable x such that $y = f(x)$ for some function f (with domain the set of all possible values of x), returns this function. To have a notation, I will write this function f as y_x. That is, y_x is the function that tells how y varies with x. We can add, multiply, and otherwise perform algebraic operations on these quantities, using $(y+z)_x = y_x + z_x$ and so on, and the usual rules of algebra apply. We can even apply functions to quantities to get new quantities: $g(y)_x = g \circ y_x$. Whenever you apply Algebra in a situation where the values of the variables are allowed to in fact vary, you cannot say that the variables stand for individual real numbers, but instead this y_x stuff is one way to formalize what is going on.

Then $\mathrm{d}y$ becomes another quantity, defined so that $(\mathrm{d}y)_x = y_x{}'$ if this exists. That is, if $y = f(x)$, then $(\mathrm{d}y)_x = f'$. While y_x tells how y varies with x, $(\mathrm{d}y)_x$ tells how *quickly* y varies with x. Since x is the identity function of itself, x_x is the identity function I, and $(\mathrm{d}x)_x$ is the derivative of I, the constant function with value 1 (which is usually also written 1). If $y = f(x)$ (so that $y_x = f$), then $(\mathrm{d}y)_x = f' = f$, while $(f'(x)\,\mathrm{d}x)_x = (f' \circ I)1 = f'$ too.

This is not enough to conclude that $\mathrm{d}y = f'(x)\,\mathrm{d}x$, because what if we use a different independent variable in place of x? Then the Chain Rule rescues us; if $y = g(u)$ and $x = h(u)$, so that $g = f \circ h$ when $y = f(x)$, then $(\mathrm{d}y)_u = g' = (f \circ h)'$, while $(f'(x)\,\mathrm{d}x)_u = (f' \circ h)h'$. The Chain Rule says precisely that these are equal.

The point of all of this is that you never need to pick an independent variable; the definitions here incorporate all possible independent variables at once. This is a common trick in abstract mathematics; if you know how a thing is supposed to work out in various contexts, then you simply define it to be a thing that works how it is supposed to. As long as the working is uniquely determined by the data given, this is a valid and precise definition.

At this point, all of the properties of differentials follow directly from the corresponding properties of derivatives. And we can recover the derivatives by dividing the differentials. As we do so, we can apply the usual rules of algebra to differentials, and it all works out!